T4 Bacteriophage Gene 32: A Structural Protein in the Replication and Recombination of DNA

by

BRUCE M. ALBERTS
LINDA FREY

Department of Biochemical Sciences,
Frick Chemical Laboratory,
Princeton University,
Princeton, New Jersey

A new type of protein essential for DNA replication and genetic recombination has been isolated from T4 bacteriophage-infected E. coli. This protein binds cooperatively to single-stranded DNA, and it catalyses DNA denaturation and recombination in physiological conditions in vitro.

GENETIC recombination involves the precise breakage and reunion of “mating” double-stranded DNA molecules at points of mutual sequence homology\(^1,2\). Recombinant DNA molecules have been shown to contain a heterozygous region, which seems to be formed during the fundamental event in the recombination process\(^3,4\). Although the actual mechanism of genetic recombination is unknown, several relatively simple models have been proposed\(^2\); these assume an unusual fluidity of DNA structure within the cell which allows efficient testing for complementarity of base pairings between strands of randomly colliding DNA molecules. For example, in the scheme proposed by Holliday\(^5\), local DNA denaturation is involved to open mating DNA helices at homologous regions, followed by DNA renaturation between single strands thereby exposed on opposite molecules. In vitro, however, the DNA double-helix is overwhelmingly stable relative to the single strands in physiological conditions\(^6\), and locally denatured regions more than a few base-pairs long should consequently occur only very rarely. This expectation is borne out by experimental studies of the stability of the short helix formed by the “cohesive ends” of isolated bacteriophage lambda DNA, for which, even in low [Na\(^+\)] (9-03 M), transient melting of twelve contiguous base-
pairs occurs with a relaxation time of about 7 days at 37°C (ref. 9). By contrast, at 37°C within the cell, the average T4 DNA molecule participates in more than one recombination exchange every 10 min, whereas it must unwind completely during each round of DNA replication.

Single-stranded regions of DNA postulated as intermediates in various models for the recombination process could easily be generated by the action of exonuclease VII,11,12 or DNA polymerase α13 within the cell, if not by denaturation. It might be expected that the complementary single-stranded regions so generated would rapidly pair by renaturation because of the overwhelming stability of the double-helically redundant DNA at 37°C. Single-stranded DNA folds on itself, however, to create imperfectly hydrogen-bonded, intra-strand helices in physiological conditions in vitro.14 These folds make the DNA bases relatively inaccessible, and thereby prevent complementary single strands from finding satisfactory pairings.14,15 As a result, raising the temperature from 37°C to 68°C increases renaturation rates as much as 1,000-fold. This is purely a kinetic effect, for the equilibrium stability of the double-helix relative to single strands is greater at the lower temperature.

We have discovered a DNA-binding protein in the T4 bacteriophage system the properties of which suggest a solution to this problem of DNA mechanics. The protein is the product of T4 gene 32. The "32-protein" is required for the genetic recombination of T4 bacteriophage DNA; in addition, it is one of several gene products known to be essential for T4 DNA replication.16

Biological Role of T4 Gene 32

The product of T4 gene 32 is required for T4 DNA replication throughout the infectious cycle: an amber mutant in this gene requires 40 min at 37°C to approximate even a single round of replication,18 whereas temperature-sensitive mutants which are allowed to begin synthesis of DNA at 25°C stop replication when shifted to a non-permissive temperature.19,20 (For one particular mutant, ts P7, all replication ceases within 1 min after a shift to 42°C: S. Riva, A. Cascino, and E. P. Geiduschek, manuscript submitted for publication; unpublished results of M. Curtis and B. M. A.) Moreover, gene dosage experiments show that gene 32 is unique among the T4 genes known to affect DNA metabolism in that its product is required stoichiometrically rather than catalytically; that is, as for structural proteins of the phage particle, the quantity of 32-protein synthesized in the newly formed virus is limited by the quantity of preformed 32-protein produced. This finding must be reconciled with the fact that about 10,000 molecules of 32-protein are made in a normal infection, and very few, if any, are used up in the construction of mature phage particles. It therefore seems that 32-protein plays a structural part in the replication of T4 DNA.

Another important biological observation concerning gene 32 is that, as first shown by Tomizawa and co-workers, its function is necessary for the formation of the hydrogen-bonded joint DNA molecules believed to be the initial products of genetic recombination.12 It should be noted that DNA replication does not seem to be required for recombination of T4 DNA, and that recombination-deficient mutants in another bacteriophage system (phage λ) replicate their DNA normally. It is therefore likely that 32-protein functions directly in both of these genetic processes.

Properties of Purified 32-Protein and its Binding to DNA

We have previously reported that at least twenty different DNA-binding proteins are synthesized after the bacteriophage infection of E. coli as judged by DNA-cellulose chromatography.23 One of the principal DNA-binding proteins was identified as the product of T4 gene 32, for it is altered after infection with bacteriophages carrying amber and temperature-sensitive mutations in this gene.23,24 The 32-protein is made in large quantities at both early and late times of infection, about 10,000 molecules accumulating per infected cell.

In the absence of a direct assay for 32-protein, the course of its purification was originally monitored by the success of acrylamide gel electrophoresis. In the work to be described stepwise elution from a single-stranded DNA-cellulose column followed by DEAE-cellulose chromatography has been used to prepare 32-protein which is electrophoretically homogeneous.

In spite of its tight binding to polyamionic DNA, 32-protein carries a net negative charge at pH 7. As estimated from a combination of sedimentation and gel filtration data, the molecular weight of the native protein is 35,000, and the axial ratio for an equivalent prolate ellipsoid is about 4 (ref. 19). Because the same molecular weight is obtained for denatured, reduced 32-protein in sodium dodecyl sulphate (SDS)-containing polyacrylamide gels, the native protein seems to consist of a single polypeptide chain.25

Purified 32-protein binds strongly to single-stranded DNA as seen by the co-sedimentation of H3-L-59 labelled protein with such DNA through stabilizing sucrose gradients. The affinity of 32-protein for DNA decreases gradually as the salt concentration is increased from 0·15 to 0·6 M, suggesting the importance of electrostatic forces in the binding. It seems likely, therefore, that the region of polypeptide chain in direct contact with the DNA includes a concentration of positively charged residues spaced so as to interact with the DNA phosphates, even though the protein as a whole carries a net negative charge.

The stoichiometry of the tight complex which 32-protein forms with single-stranded DNA at low salt concentrations has been examined by sucrose gradient sedimentation of a fixed quantity of labelled 32-protein in the presence of varying amounts of the circular, single-stranded DNA from bacteriophage fd.26 At lower concentrations of DNA, two distinct peaks of radioactive protein are seen; one sediments rapidly with the DNA, the second at the slow rate characteristic of the free protein. The free protein peak is absent above a weight ratio of DNA to protein of 1 : 12. The complex therefore contains about one protein molecule of 35,000 molecular weight for every ten single-stranded DNA nucleotides. Because of the small size of the nucleotides at most 70 Å, whereas 32-protein may be as much as 120 Å long, adjacent molecules of 32-protein could overlap in the complex. Consistent with this expectation, it was shown previously that in crude extracts 32-protein binds cooperatively to single-stranded DNA-cellulose.

To determine whether the purified 32-protein also binds cooperatively to DNA, two different concentrations of 32-protein (containing the same amount of tritiated 32-protein) were mixed with a constant amount (large excess) of fd DNA at an elevated salt concentration where the complex is otherwise marginally stable. The results of sucrose gradient sedimentation analyses are shown in Fig. 1. It is clear that a 14-fold increase in 32-protein concentration dramatically increases its DNA affinity. This result requires that 32-protein molecules interact with each other in the complex (see caption to Fig. 1), in which case there must be at least eight times greater than its affinity for an isolated site, for strong cooperative binding is observed even in conditions where the number of isolated
sites available exceeds the number of contiguous sites by at least this factor (Fig. 1).

The highly cooperative nature of the DNA affinity of 32-protein should cause it to bind to DNA in long clusters even in conditions of large DNA excess. Direct evidence for such clustered binding is obtained when labelled 32-protein is mixed with a large excess of fd DNA and sedimented through sucrose gradients at low salt concentrations. In this experiment, a larger portion of the 32-protein sediments ahead of the main DNA peak, being tightly bound to a small fraction of the DNA molecules. The mean size of a 32-protein cluster in these conditions must therefore be an appreciable fraction of the length of an fd DNA molecule (6,600 nucleotides)22. Clustered binding to poly dA can also be detected by this method, indicating that the cooperativity observed is the result of direct stabilizing interactions between adjacent 32-protein monomers. This view is also supported by our finding (unpublished) that 32-protein self-aggregates in the absence of DNA at a concentration of 0.5 mg/ml or higher.

![Fig. 1. Cooperative binding of 32-protein to single-stranded DNA. Purified fd DNA (10 µg) was mixed with about 0.5 µg (a) or 7 µg (b) of 3H-labelled 32-protein in 0.2 ml of 0.02 M Tris-HCl (pH 8)–0.2 M Na₂EDTA–0.3 M NaCl–100 µg/ml bovine serum albumin (BSA)–10% glycerol–1 mM β-mercaptoethanol at 4°C. After 20 min, the mixture was layered at 4°C onto a 5 ml, 0–30 per cent sucrose gradient prepared in the same buffer. Following centrifugation for 2 h at 46,000 r.p.m. in the Spinco SW 50 rotor, 0.15 ml fractions were collected and monitored for radioactivity by standard techniques. Recoveries of 3H-protein added averaged about 75 per cent. Concentrations of 32-protein (A₂₆₀=11 µg/ml) and fd DNA (A₂₆₀=23.8 µg/ml) in 0.15 M NaCl–0.015 M sodium citrate, pH 7.0 were determined by absorbance measurements. Note that, for ordinary binding, the protein distribution would have been identical in the two above experiments, for (free protein) (bound protein) = K₀ (free DNA sites), and the concentration of free DNA sites was held essentially constant.](https://example.com/figure1)

Although 32-protein binds very tightly to all single-stranded DNAs tested, including the synthetic polynucleotide poly dA, no binding of the purified protein to double-stranded DNAs or to H17 RNA could be detected by sucrose gradient sedimentation at 4°C.

**Denaturation of DNA with 32-Protein**

Histones and polyanimes bind to the double-helical form of DNA more tightly than to single strands and thereby raise the temperature required for DNA denaturation22. Conversely, the strong selective affinity of 32-protein for single-stranded DNA should lower the thermal denaturation temperature of double-stranded DNA. A precedent for such an effect is the destabilization of DNA observed in the presence of pancreatic ribonuclease which likewise preferentially binds to DNA single strands23.

Because single-stranded DNA is fully hyperchromic when complexed with 32-protein, any denaturation of double-stranded DNA which occurs in its presence should be accompanied by the large hyperchromic change that is characteristic of this helix-coil transition. By this criterion, double-stranded T4 DNA is not denatured in the presence of excess 32-protein in a variety of ionic conditions at temperatures up to 37°C. By contrast, poly dAT, which normally has a Tm about 10°C lower than T4 DNA (35°C as against 41°C in 0.01 M KCl–0.01 M MgSO₄), is readily denatured by 32-protein even at 25°C. Typical kinetics for poly dAT denaturation by 32-protein are shown in Fig. 2. In the presence of 0.01 M Mg²⁺, half-denaturation of poly dAT by 32-protein is attained in about 20 min in the conditions used; this denaturation is reversible, for the absorbance at 260 nm can be restored to its original value either by addition of NaCl to 0.5 M at 25°C (to dissociate 32-protein), or by direct cooling to 4°C. The initial rate of denaturation at 25°C is reduced with increasing [Mg²⁺], decreasing at least 15-fold when [Mg²⁺] is increased from 0.01 M to 0.04 M, and increasing about three-fold when all Mg²⁺ is removed.

To denature poly dAT at 25°C, ΔG for the coil–complex reaction (single-stranded DNA coil + 32-protein → DNA – protein complex) must be sufficiently negative to make ΔGΔt < 0 at that temperature. The ΔGΔt < 0 should be about 1.5 kcal/mole base-pair for poly dAT in 0.01 M Mg²⁺ at 25°C (Tm = 65°C) (ref. 28). Consequently, to obtain denaturation, ΔGΔt should be as large as possible, or the concentration of binding sites (ten single-stranded DNA nucleotides) because Kₐₚₐₜₜ = exp (–ΔG/RT) (free protein) (free sites) for both denaturation of poly dAT in 0.01 M Mg²⁺ at 25°C with 170 µg/ml of 32-protein (Fig. 2) will require an effective dissociation constant for 32-protein of < 1.1 × 10⁻¹⁰ M. This is in agreement with direct measurements in sucrose gradients in similar conditions, which yield a dissociation constant (averaged for cooperativity) of < 10⁻⁸ M for the 32-protein complex with fd DNA.

![Fig. 2. Poly dAT denaturation catalysed by 32-protein at 25°C. In addition to the concentration of MgSO₄ indicated, each sample contained 10 µg/ml of poly dAT (from A. Kornerup, 0.02 M NaOH–0.01 M NaCl), 170 µg/ml 32-protein, 0.01 M KCl, 5 mM Tris-HCl (pH 8.1), 1 mM β-mercaptoethanol, 0.1 mM Na₂EDTA, and 2 per cent glycerol. Samples were placed in cuvettes in the thermostated compartment of a (U)rafft spectrophotometer, so that the absorbance could be monitored automatically. At time zero, concentrated 32-protein was added to start the reaction. Similar results have been obtained at a KC1 concentration of 0.12 M.](https://example.com/figure2)

**Renaturation of DNA with 32-Protein**

As already noted, the renaturation of purified DNA in physiological conditions in vitro is an extremely slow process, because of the intrastrand folding in denatured DNA. Our results imply, however, that in the T4 system single-stranded DNA does not fold as much in vivo, but is instead always present as a tight complex with 32-
protein. We find that if DNA saturated with 32-protein sediments only about 1-3 times faster than the free DNA, although its mass is thirteen times greater. This means that the frictional coefficient of 32 DNA increases about six-fold in the complex. Because frictional coefficients only double when single-stranded DNA is unfolded in alkali\(^2\), complexed DNA must be held in a highly expanded conformation by 32-protein. (This expansion can be seen directly by electron microscopy; personal communication of H. Delius.) DNA in such a conformation might be expected to renature much more rapidly than free de

denatured DNA at low temperatures.

The rate of renaturation of DNA covered with 32-protein was measured by an absorbance assay similar to that used to monitor denaturation in Fig. 2. In this case, a decrease in the absorbance of single-stranded DNA is expected proportional to the amount of reformed double-helix. Typical results obtained for single-stranded T4 DNA in the presence of excess 32-protein at 25°C are presented in Fig. 3. It can be seen that a rapid decrease in absorbance is observed in 0-04 M Mg\(^{2+}\), representing a more than 1,000-fold acceleration of the renaturation rate without 32-protein. The dependence of this reaction on [Mg\(^{2+}\)] is the reverse of that found for denaturation in Fig. 2: the renaturation rate drops about four-fold in 0-01 M Mg\(^{2+}\), whereas no renaturation is detected without Mg\(^{2+}\).

A more sensitive measure of the course of renaturation is obtained from CsCl gradients, where renatured DNA has a lower buoyant density than single strands\(^1\). This assay can be used for kinetic analyses, for the addition of concentrated CsCl dissociates 32-protein from the DNA and prevents further renaturation. By this technique, it has been found that the rate of renaturation of T4 DNA in the presence of excess 32-protein is proportional to the square of the DNA concentration, showing that, as in the uncatalysed reaction, the rate measured is that for the nucleation of complementary pairings. Some second order rate constants for the renaturation of T4 DNA single strands of molecular weight \(5 \times 10^9\) are listed in Table 1. Note that the rate of renaturation catalysed by 32-protein at 37°C can exceed the uncatalysed rate observed in standard conditions (68°C in 1 M NaCl).

If 32-protein accelerates DNA renaturation by imparting favourable conformation to the single strands, the bases of which would otherwise be inaccessible on highly folded chains, the dependence of renaturation rate on the ratio of 32-protein to DNA should be unusual. Below the saturating protein : DNA ratio of 12 : 1, the rate of renaturation should be drastically lowered, initial rates being proportional to at least the square of the amount of 32-protein added. Results of renaturation assays performed at sub-saturating 32-protein levels are shown in Fig. 4, where it is seen that a four-fold drop in the concentration of 32-protein decreases the rate of renaturation of T4 DNA at least twenty-five-fold, as expected. A second expectation is that above a protein : DNA ratio of 12 : 1, additional 32-protein should not further increase renaturation rates. This prediction has also been confirmed (experiment not shown).

On the basis of these results, we conclude that 32-protein accelerates renaturation in physiological conditions by forcing DNA single strands into an unfolded conformation which leaves their bases available for pairing during chance collisions between complementary strands. It seems likely that the bound 32-protein is rapidly displaced from the rewinding single strands as the double-helix forms (see Fig. 3).

![Graph showing DNA renaturation catalysed by 32-protein at 25°C.](image)

**Fig. 3.** DNA renaturation catalysed by 32-protein at 25°C. Single-stranded T4 DNA was prepared by alkaline denaturation and dialysis into low salt buffer as described by Studier\(^2\), except that the DNA was sheared in alkali to molecular weight \(6 \times 10^9\) before the dialysis. The DNA serves as a control, for it is not self-complementary and therefore cannot renature. At zero time, 0.00 g/ml of each DNA was mixed in separate cuvettes with 170 μg/ml of 32-protein in the buffer used in Fig. 2. At the times indicated, MgSO\(_4\) was added in a 1:1 stock to both 32 and T4 DNA reactions. A decrease in absorbance of about 0.072 units is expected for full renaturation of the T4 DNA.

**Table 1. Renaturation rates for T4 DNA single strands of molecular weight \(5 \times 10^9\).**

<table>
<thead>
<tr>
<th>32-Protein concentration (μg/ml)</th>
<th>Ionic composition</th>
<th>Temperature (°C)</th>
<th>(K(_r)/1. mole(^{-1}).s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.01 M NaCl</td>
<td>37</td>
<td>22</td>
</tr>
<tr>
<td>10</td>
<td>0.04 M MgSO(_4)</td>
<td>37</td>
<td>300</td>
</tr>
<tr>
<td>10</td>
<td>0.04 M MgCl(_2)</td>
<td>37</td>
<td>300</td>
</tr>
<tr>
<td>10</td>
<td>0.01 M MgCl(_2)</td>
<td>37</td>
<td>23</td>
</tr>
<tr>
<td>10</td>
<td>0.01 M NaCl</td>
<td>37</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>0.01 M MgSO(_4)</td>
<td>37</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>0.01 M MgCl(_2)</td>
<td>37</td>
<td>30</td>
</tr>
</tbody>
</table>

For the 32-protein catalysed reaction: (1) renaturation rates decrease with storage of the protein, so that the maximum rates are probably greater than those listed; (2) between 0.01 M and 0.04 M Mg\(^{2+}\), the rate at 37°C is roughly proportional to [Mg\(^{2+}\)]; (3) the efficiency of renaturation is reduced to levels as high as 10. After 90 min, the reaction was quenched by addition of concentrated CsCl (to \(1=1000\)) plus 50 μg of sodium dodecyl sulphate. The bands of DNA were polymers which were then photographed after 20 h at 44,000 r.p.m., in the Spinco model B analytical ultracentrifuge. The same model has shifted to a buoyant density representing molecules which are about two-thirds double-helical, as expected from the pairing of strands randomly cut from a longer sequence\(^2\). As controls, no density shift was observed when 32 DNA was processed in an identical manner, whereas fragments of single-stranded T7 DNA renatured rapidly in the presence of 32-protein.
Function of 32-Protein in Genetic Recombination

The denaturation of poly dAT by 32-protein at 25°C indicates that within the cell, the cooperative action of 32-protein should generate fluctuating regions of local denaturation in T4 DNA. T4 DNA, with or without added tRNA, in early steps of genetic recombination can be explained by this ability to open up local regions of native DNA, while simultaneously facilitating helix formation between matching, complexed single strands. Both of these functions are probably necessary for efficient testing of complementary pairings between double-stranded DNA molecules. In addition, experiments with infected cells have suggested that the formation of single-strand breaks ("nicks") in double-stranded intracellular T4 DNA is a prerequisite for the initiation of strand exchanges. Order-of-magnitude calculations, based on the treatment of DNA renaturation formulated by Wetmur and Davidson, suggest that random collisions between specifically nicked, double-stranded T4 DNA molecules (see ref. 5) might be efficient enough in the presence of 32-protein to account for the high recombination rates observed in the T4 system (unpublished results). This, however, merely indicates the mechanism of genetic recombination may be relatively simple, for the properties of 32-protein are also compatible with most other types of models suggested for the recombination process.

Function of 32-Protein in DNA Replication

The genetic results which indicate a structural role for 32-protein in DNA replication suggest that local unwinding by 32-protein might be required in the replication fork in order for productive replication to proceed. In agreement with this role, in vitro experiments have revealed that T4 DNA polymerase uses single-stranded DNA templates much more rapidly in the presence of 32-protein than in its absence (unpublished results of Huberman, Kornberg and B. M. A.). This stimulation is probably the result of favorable template alignment by 32-protein. Because stimulation is not observed in similar experiments with E. coli DNA polymerase, a direct interaction of T4 polymerase with 32-protein may also be involved.

In a normal T4 bacteriophage infection, the number of replication forks present per cell increases linearly until 30 min post-infection (29°C). The rate at which the number of polymerization observed at each fork is constant throughout this period, this rate must be independent of the level of 32-protein, inasmuch as this increases continuously during infection. Yet the gene dosage experiments, which reveal a direct proportionality between the quantity of phage DNA produced and the quantity of 32-protein present, seem to demand that the overall rate of DNA synthesis be proportional to the amount of 32-protein made. To account for these facts, we propose that a functioning replication fork has a unique tertiary structure that contains a fixed number of 32-protein molecules. About sixty new replication forks are eventually generated in a normal T4-infected cell, so that each fork could incorporate no more than 170 molecules of 32-protein.

If each cycle of DNA replication begins at a special point on the T4 genome, new replication forks must be generated only at a unique nucleotide sequence. Both during this process and as the replication fork travels, 32-protein may interact with other proteins in addition to T4 DNA. Likely candidates for such proteins include the products of T4 genes 41, 44, 45, 59 and 62, all of which have as yet unidentified functions essential for T4 DNA replication. Further studies involving 32-protein may therefore provide a fresh insight concerning the unknown mechanism by which DNA is replicated in biological systems.

Preparation of Homogeneous 32-Protein

An E. coli culture, grown to 5 x 10^10 cells per ml at 32°C in M-9 minimal media containing 0.3 per cent casein hydrolysate and 0.3 per cent glucose was incubated at 30°C and stored at 20°C. Cells (50 g) were broken by lysis in 100 ml of 0.005 M Tris-HCl (pH 8.1) — 0.001 M MgCl_2 - 2 mM CaCl_2 - 1 mM β-mercaptoethanol - 1 mM Na_2EDTA containing 20 μg/ml pancreatic deoxyribonuclease I. The slurry was centrifuged at 27,000 x g, washed twice with 200 ml of 0.01 M Tris-HCl (pH 8.1) — 0.01 M NaCl — 5 mM Na_2EDTA — 1 mM β-mercaptoethanol (buffer A) to remove the divalent cations necessary for the activity of deoxyribonuclease I. After centrifugation to remove a light precipitate, the dialysed extract was made 10 per cent glycerol and forced at 100 ml/h through a column containing 20 ml packed volume of denatured calf thymus DNA-cellulose (approximately 1 mg of DNA per ml). The DNA-cellulose (7 cm x 3 cm) had been equilibrated with a buffer consisting of 10 per cent glycerol in buffer A, and this basic buffer was used for an 80 ml rinse and for elutions in which increasing concentrations of NaCl were added. The column was eluted at 20 ml/h, in 40 ml steps of 0.15, 0.40, 0.60, 0.90, 2.0 and 5 M NaCl. The peak 2.0 M NaCl eluting fraction contained 32-protein as its principal component. This fraction (8 ml) was dialysed against 0.01 M Tris-HCl (pH 8.1) — 10 per cent glycerol — 1 mM Na_2EDTA — 1 mM β-mercaptoethanol (buffer B) and applied to a 7 cm x 0.8 cm column of DEAE-cellulose (Whatman DE32). The column was washed with 5 ml of buffer B and then eluted with a 30 ml linear gradient of 0.0-0.5 M NaCl in this buffer. Fractions of 1.3 ml were collected every 20 min. The 32-protein (A_280/A_260 absorbance ratio of 1.7) appeared in three adjacent fractions with a mean NaCl concentration of 0.20 M. These fractions were used directly for the studies to be described, or concentrated further by vacuum dialysis against buffer B containing 0.05 M KCl. Approximately 8 mg of electrophoretically homogeneous 32-protein is obtained by this procedure. As determined by the subsequent recovery of purified "H-labelled 32-protein added to crude extracts, this represents about a 65 per cent yield. An identical procedure was used on a smaller scale for preparation of "H-labelled 32-protein, except that the cells were grown at 25°C and labelled with 500 μCi of "H-lucine after 35 min of infection. Unless otherwise stated, all operations were carried out at 4°C; at this temperature, concentrations of 32-protein (>200 μg/ml) may be kept for several weeks. At ~80°C, 32-protein has been stored for up to 8 months without a noticeable change in its DNA affinity.

This work was supported by grants from the US National Institutes of Health and the American Cancer Society, and in particular by Eugene Higgins’ Trust Fund, and facilities made available by the Whitehall and John A. Hartford Foundation to the Department of Biology, Princeton University. We thank Walter Kauzmann, Noboru Suseoka, and Jacques Fresco for discussions, and Barbara Banman for skillfully performing the analytical ultracentrifugations.

Received June 11; revised July 7, 1970.

Three Variable-Gene Pools common to IgM, IgG and IgA Immunoglobulins

by

HEINZ KÖHLER
AKIRA SHIMIZU
CLAUDINE PAUL
VIRGINIA MOORE
FRANK W. PUTNAM

Department of Zoology, Indiana University, Bloomington, Indiana 47401

Immunoglobulins can be divided into three principal classes according to their different antigenic and functional properties. The class differences reside in the heavy chains, whereas the light chains, which are chemically and immunologically defined as \( \alpha \) or \( \lambda \) chains, are shared by all three immunoglobulin classes. The IgA immunoglobulin class is characterized by the \( \gamma \) chain, IgG by the \( \gamma \) chain and IgM by the \( \mu \) chain. In spite of these structural and functional differences certain common features exist for the three chief classes. A common evolution for all immunoglobulins is evident from comparison of the chemical structure of heavy and light chains from different species. Furthermore, there is a general cooperative and sequential relationship in the immune response of all three immunoglobulin classes. Finally, serological cross-reactions between IgM and IgG have been observed by a number of workers. In fact, the earliest evidence for a heavy chain gene was the finding by Todd of a common allotypic marker in rabbit IgG and IgM. Thus the question arises where in the primary structure of the \( \alpha \), \( \gamma \) and \( \mu \) chains the differences and similarities of the three principal immunoglobulin classes are localized. Earlier we made the first comparison of an extended sequence of a \( \mu \) chain (Oui) with a \( \gamma \) chain (Daw) that showed that in the first 105NH\(_4\)-terminal residues of these two human heavy chains of different classes the homology in sequence was almost 75 per cent. Subsequently, we undertook sequence analysis of several human IgM Waldenström macroglobulins and IgA myeloma proteins to ascertain if this high degree of homology of two different heavy chain classes was merely accidental. We now report additional heavy chain sequences from the NH\(_4\)-terminus of four IgM macroglobulins and one IgA myeloma globulin. Comparison of these data with the NH\(_4\)-terminal sequences of heavy chains from other laboratories has revealed homologies of different degrees which suggest the existence of two new variable heavy chain subgroups, V\(_{HIII}\) and V\(_{HIV}\) in addition to the V\(_{HI}\) and V\(_{HII}\) subgroups already proposed.

Comparison of the NH\(_4\)-terminal region of several \( \alpha \), \( \gamma \) and \( \mu \) chains indicates that four variable-sequence subgroups are common to heavy chains. The existence of three independent variable-gene pools common to the three major immunoglobulin classes is confirmed.

Determination of Amino-acid Sequence

Because most of the heavy chains in immunoglobulins have a blocked NH\(_4\)-terminal residue which is a cyclic d form of glutamine (pyrrolidone carboxylic acid or PCA)\(^{10}\), a specific method to isolate blocked NH\(_4\)-terminal peptides from the whole molecule can be used\(^{11}\). IgM protein (1-2 moles) was digested with staphilin for 2 h at 37°C and pH 8. The soluble digest was then applied to a Dowex 50 × 2 cm column previously washed with 1 M HCl and distilled water until neutral. The first peak eluted contained glycopeptides and the second peak a tetrapeptide which was ninhydrin-negative on paper. The amino-acid composition of the tetrapeptide from proteins Ou, Di and Re varied in only one residue. In all three cases a 10 min and 120 min incubation with carboxypeptidase A released two residues in different yields. By this method the amino-acids at positions three and four were established. Because the first residue could be assumed to be PCA, the residue in position two could be deduced. In another experiment, we determined the partial sequence of the nineteen NH\(_4\)-terminal residues of protein Di by study of tryptic peptides from the NH\(_4\)-terminal fragment obtained by CNBr cleavage.

Attempts to isolate a blocked peptide failed with the IgM proteins W0 (personal communication from A. van Dalen) and Na and with the IgA protein Ha. But we found a free NH\(_4\)-terminus on these heavy chains by means of the denaturation technique\(^{12}\). The NH\(_4\)-terminal residue in the unblocked heavy chains of proteins W0, Na and Ha was glutamic acid. These results confirm earlier data from this laboratory which showed that in about half of the \( \mu \) chains studied the NH\(_4\)-terminus was not blocked\(^{13}\). The \( \mu \) chains of IgM Wo and Na and the \( \alpha \) chain of IgA Ha were subjected to the automatic Edman degradation method\(^{12}\) using the Beckman sequencer model 890. Identification of the degraded PTH-amino-acid was done by thin-layer chromatography\(^{14}\) and by gas chromatography using a combination of three different columns. In one